

<<视觉语音情感识别>>

图书基本信息

书名：<<视觉语音情感识别>>

13位ISBN编号：9787030373212

10位ISBN编号：7030373219

出版时间：2013-4

出版时间：科学出版社

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<视觉语音情感识别>>

前言

计算机科学技术的迅速发展改变了人们的工作和生活方式。

人机交互(human computer interaction, HCI)技术是突破计算机与人类交互瓶颈的重要技术,它有力地促进了计算机的普及应用。

为了使人类与计算机之间进行更加智能、更加自然的交互,新型的人机交互技术正在逐渐成为研究热点。

人们不仅希望能以更方便、更高效、更自然的方式操纵计算机,而且还希望计算机能理解人的情感,并提供有价值的增值服务。

因此,情感识别已经成为人机交互领域亟待突破的关键技术之一。

所谓情感识别,就是利用计算机分析特定表演者的面部表情、姿势和语音信号及其变化过程,进而确定该表演者的内心情绪或思想活动,实现人机之间更智能、更自然的交互。

情感分析在许多领域都有着潜在的应用价值,如心理学研究、图像理解、脸部动画合成、视频检索、机器人技术和虚拟现实技术等领域。

随着多媒体技术的发展,基于音视频的情感分析、识别研究对增强计算机的智能化和人性化、开发新型人机环境,以及推动多媒体技术和信号处理等相关领域的发展有着重要的意义。

本书作者近几年来在国家自然科学基金项目(60673190、61003183、61272211)的支持下,开展了基于视频和音频的情感识别技术的研究,包括视频和音频的情感特征的提取、适合于音视频情感分析、分类和识别的算法或分类器、多分类器融合的情感识别方法以及情感分析、识别的应用等。

本书是在这些项目的研究成果基础上,系统化地加以归类总结撰写而成的。

全书内容分为8章。

第1章为绪论,简要地介绍了视觉语音情感识别产生的背景和意义、研究内容及其应用领域。

第2章为视觉语音情感识别技术概况,包括情感识别的框架、情感的定义与分类、情感信号的预处理、常见的情感特征提取与选择、现有的情感识别方法,以及情感识别的难点及新动向。

第3章为基于视觉信息的情感特征提取方法,内容包括基于视频、图像的表情特征提取的相关方法,这些方法是对已经提出的一些方法的改进,并提出自己所设计的方法。

第4章为基于视觉信息的情感识别方法,内容包括了较有特色的相关表情分类、识别方法,如针对样本的不均衡性,提出并实现的最大间隔最小体积球形支持向量机的表情识别方法;针对细微表情的识别问题,提出并实现的混合特征结合分类树的细微表情识别算法;针对视频信息存在噪声和部分遮挡问题,提出并实现的基于模糊深隐马尔可夫模型的图像序列表情识别方法。

第5章为语音情感特征选择提取方法,内容包括了语音情感测试库的构建、个性化和基于多重分形理论的语音情感特征提取方法、语音情感特征有效选择方法,以及基于流形学习的语音情感特征降维等方法。

第6章为语音情感识别方法,内容包括了基于选择性特征的决策树的语音情感识别方法和基于改进有向无环图的分层语音情感识别方法。

第7章为视觉语音融合情感识别方法,内容包括基于D-S证据理论的多粒度语段融合情感识别方法、多分类器融合方法和具有噪声过滤功能的分类器协同训练半监督主动学习方法等。

第8章为情感分析的应用,内容包括基于表情动作单元参数的逼真表情动画方法和E-learning环境中的情感分析应用方法。

本书较全面地总结了课题组近年来的有关视觉语音情感识别的研究成果,内容系统、深入浅出、方法有新颖性和创新性。

适合从事音视频信息的处理,特别是从事智能人机交互、计算机视觉与听觉、模式识别与人工智能以及智能动画等领域的科技工作者阅读参考。

本书也可作为情感计算、新型人机交互和智能信息处理等课程的研究生教材。

本书由詹永照总体负责,毛启容负责组织实施。

本书的第1、2章由毛启容和成科扬撰写,第3章由林庆撰写,第5、6章由毛启容撰写,詹永照负责了其余各章的撰写和全书的统稿工作。

<<视觉语音情感识别>>

完成这些项目的老师和合作者还有张建明、文传军、叶敬福、曹鹏、周庚涛、陈亚必、刘娟、李婷、陆捷荣、张娟、胡敏灵、徐莉婷、刘云、孔建等，他们在课题研究中刻苦钻研，做了有意义的探索性工作，也为本书的完成做出了重要贡献。

在此谨向他们表示衷心的感谢。

本书的撰写过程也参考了国内外研究者的研究成果和资料，也一并向他们致谢。

情感计算属于交叉学科的新兴研究领域，由于我们的水平有限，书中难免有疏漏之处，敬请读者不吝指正。

詹永照 2012年10月于江苏大学

<<视觉语音情感识别>>

内容概要

《视觉语音情感识别》作者近几年来在国家自然科学基金项目（60673190、61003183、61272211）的支持下，开展了基于视频和音频的情感识别技术的研究，包括视频和音频的情感特征的提取、适合于音视频情感分析、分类和识别的算法或分类器、多分类器融合的情感识别方法以及情感分析、识别的应用等。

《视觉语音情感识别》是在这些项目的研究成果基础上，系统化地加以归类总结撰写而成的。

全书内容分为8章。

《视觉语音情感识别》较全面地总结了课题组近年来的有关视觉语音情感识别的研究成果，内容系统、深入浅出、方法有新颖性和创新性。

<<视觉语音情感识别>>

书籍目录

前言 第1章绪论 1.1视觉语音情感识别的产生背景 1.2视觉语音情感分析的研究内容 1.3视觉语音情感识别的应用领域 1.3.1表情识别的应用 1.3.2语音情感识别的应用 参考文献 第2章视觉语音情感识别技术概况 2.1情感识别框架 2.2情感描述模型 2.2.1情感的定义 2.2.2情感的分类 2.3视觉语音信号预处理 2.3.1人脸表情图像预处理 2.3.2情感语音信号预处理 2.4情感特征提取 2.4.1视觉信息情感特征提取 2.4.2语音情感特征提取 2.5情感特征选择 2.6常用的情感识别模型 2.6.1基于相似性的情感识别模型 2.6.2基于连接机制的情感识别模型 2.6.3基于概率模型的情感识别模型 2.6.4基于集成学习的情感识别模型 2.7视觉语音情感识别的挑战 2.8视觉语音情感识别的新动向 参考文献 第3章基于视觉信息的情感特征提取方法 3.1概述 3.2基于小波分解和优选VLBP的表情特征提取方法 3.2.1表情图像的小波分解 3.2.2小波分解图像的情感特征提取 3.2.3实验结果与分析 3.3基于多频域LBP—TOP的人脸表情特征提取方法 3.3.1 LBP—TOP算子 3.3.2多频率图像分块LBP—TOP特征提取 3.3.3实验结果与分析 3.4基于VLBP与光流的混合情感特征提取 3.4.1眼睛区域的小波分解分块VLBP特征提取 3.4.2特征点自动标注的嘴部光流特征提取 3.4.3基于混合特征的表情识别 3.4.4实验结果与分析 3.5基于Gabor变换的表情图像特征提取方法 3.5.1小波变换与多分辨率分析 3.5.2 Gabor变换 3.5.3人脸表情图像的网格化 3.5.4基于Gabor小波变换的表情弹性图的构造 3.5.5实验结果与分析 3.6基于积分图像的表情特征提取方法 3.6.1积分图像的概念 3.6.2积分图像表情特征提取 3.6.3实验结果与分析 3.7一种加权矩形提取表情特征的方法 3.7.1矩形模板设计 3.7.2基于加权矩形的表情特征提取 3.7.3实验结果与分析 3.8本章小结 参考文献 第4章基于视觉信息的情感识别方法 4.1概述 4.2最大间隔最小体积球形支持向量机 4.2.1最大间隔球形支持向量机 4.2.2最大间隔最小体积球形支持向量机 4.2.3模型性能分析 4.2.4基于最大间隔最小体积球形支持向量机的表情识别 4.3混合特征结合分类树的细微表情识别算法 4.3.1混合特征提取 4.3.2基于分类树的表情识别 4.3.3实验结果与分析 4.4基于模糊深隐马尔可夫模型的图像序列表情识别方法 4.4.1模糊深隐马尔可夫模型 4.4.2模糊深隐马尔可夫模型的特性 4.4.3基于模糊深隐马尔可夫模型的图像序列表情识别 4.4.4实验结果与分析 4.5本章小结 参考文献 第5章语音情感特征选择提取方法 5.1概述 5.2情感语音库的录制 5.2.1语音情感数据库 5.2.2音视频情感数据库 5.2.3音视频情感的有效性分析 5.3语音情感特征提取方法 5.3.1传统声学语音情感特征分析与提取 5.3.2基于导数的非个性化语音情感特征提取方法 5.3.3基于多重分形理论的语音情感特征提取方法 5.4语音情感特征选择方法 5.4.1基于神经网络贡献分析的语音情感特征选择 5.4.2基于遗传算法的语音情感特征选择和分类器参数优化方法 5.4.3基于类集 / 类对的两级语音情感特征选择方法 5.4.4基于扩展测地距离的语音情感特征有效性分析方法 5.4.5实验结果与分析 5.5基于流形学习的语音情感特征降维 5.5.1 Isomap算法描述 5.5.2 ELE算法描述 5.5.3基于增量流形学习的情感特征降维方法 5.5.4实验结果比较与分析 第6章语音情感识别方法 第7章视觉语音融合情感识别方法 第8章情感分析的应用 参考文献

<<视觉语音情感识别>>

章节摘录

版权页：插图：2.4.2语音情感特征提取 总体上讲，人类的语音传达了两类信息。

一类是语言学信息，它以某一种语言规范准确（或者近乎准确）地确定了说话人的定性目的。

另一类是超语言学信息，它是以实现说话人定性目的方式的细微变化来传达的。

超语言学信息不能传达语义信息，但它能传达有关说话人当前情绪和态度的信息。

此外，它或许还能传达有关说话人口音和社会阶层特征等信息。

超语言学信息包括韵律学相关的基频和响度的变化，以及语音质量相关的声音频谱上的变化，只要这些变化不会引起语言学信息的失真。

目前，在语音情感特征提取领域中，情感特征可以粗略地分为基于超语言学的语音情感特征和基于语言学的语音情感特征（包括词汇、句法等）。

1.基于超语言学的语音情感特征 基于超语言学的语音情感特征可分为三大类：韵律学特征、音质特征和频谱特征。

韵律学特征主要包括音调、音强和语速等。

韵律学特征主要包括如下三类。

（1）基频相关特征：主要是指基音频率及其均值、变化范围、变化率与均方差等特征。

（2）能量相关特征：主要是指短时平均能量、短时能量变化率、短时平均振幅、振幅平均变化率及短时最大振幅等。

（3）时长相关特征：主要指语速、短时平均过零率等。

音质特征主要有呼吸声、明亮度特征（低频能量和高频能量的比值，用以反映语音的清亮特性）、共振峰和喉化音等；频谱特征主要包括Mel倒谱系数（Mel frequency cepstrum coefficients, MFCC）及其衍生参数等。

此外，基于这三类语音情感特征的不同语段长度的统计特征是目前使用最普遍的特征参数之一，如特征的平均值、变化率及变化范围等。

传统的基于声学的语音情感特征在整个语音情感特征中起着比较重要的作用，是使用最多的语音情感特征。

但是这类情感特征中包含了说话者与生俱来的说话特征，如语速快、嗓门儿大等特点，也包含了说话内容的信息，使这类特征中的绝大部分特征参数的分布会随着说话者和说话内容的变化而产生较大的变化，从而使当待识别语音的说话者不在训练语音情感库中，说话内容发生变化时，识别率急剧下降。

除了上述的基于声学的传统语音情感特征外，近年来，国内外研究者提出了新的基于声学的情感特征以及情感特征使用的新方法。

<<视觉语音情感识别>>

编辑推荐

《视觉语音情感识别》适合从事音视频信息的处理，特别是从事智能人机交互、计算机视觉与听觉、模式识别与人工智能以及智能动画等领域的科技工作者阅读参考。

《视觉语音情感识别》也可作为情感计算、新型人机交互和智能信息处理等课程的研究生教材。

<<视觉语音情感识别>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>